

Intelligence Artificielle pour le Génie Logiciel

Extraction de métriques et utilisation d’algorithmes prédictifs pour la gestion de projet

Quentin Perez, Christelle Urtado et Sylvain Vauttier

7 novembre 2022

Objectif de ce TP

L’objectif de ce TP est de vous montrer un exemple d’utilisation de méthodes d’intelligence artificielle dans le domaine du génie logiciel. Nous allons utiliser l’apprentissage supervisé afin de prédire les développeurs dits ”expérimentés” dans un projet industriel disponible sur GitHub. Nous allons également créer un prédicteur simple du nombre de développeurs expérimentés par rapport à la taille du projet. Les données que nous utiliserons sont des métriques logicielles. Nous allons aussi traiter d’autres questions simples mais néanmoins représentatives des possibilités qu’offrent la data-science, l’IA et la statistiques en GL.

Environnement de travail

Dans ce TP, nous allons travailler avec un environnement Python 3 et avec des *notebooks* Jupyter. Jupyter permet de créer des *notebooks* exécutables via un navigateur web ou des IDE le supportant. L’environnement Jupyter permet l’exécution, la visualisation des résultats et leur sauvegarde. Un carnet peut ainsi être exécuté et sauvegardé avec ses résultats (graphiques ou non) sans avoir à ré-exécuter l’application pour les obtenir. Nous allons également utiliser différents *packages* Python pour :

- Le clonage de projet et la manipulation de données Git, notamment les tags à l’aide de [GitPython](#) ;
- L’ouverture de fichiers CSV et la manipulation de données avec [Pandas](#) ;
- L’apprentissage machine avec la bibliothèque [Scikit-Learn](#) ;
- La visualisation avec [Matplotlib](#).

Environnement d’exécution

Pour exécuter le notebook Jupyter servant de support à ce TP nous allons utiliser Google Colab. Il fournit un environnement virtuel standard type VM vous permettant de ne pas dépendre d’une installation de l’ensemble de l’environnement sur votre machine locale (qui est lourd et parfois mystérieux si ce n’est mystique...). Le Colab que nous allons utiliser pour ce TP est accessible à l’aide de l’URL suivante :

<https://colab.research.google.com/drive/1dkdA1rcTDjyyMPj07ICgyVOOND-kWjPv?usp=sharing>

Rendu

Ce TP sera à rendre sur un dépôt Moodle qui sera ouvert prochainement.